Chemistry & Biology, Vol. 10, 713–722, August, 2003, ©2003 Elsevier Science Ltd. All rights reserved. DOI 10.1016/S1074-5521(03)00172-8

A Synthetic Miniprotein that Binds Specific DNA Sequences by Contacting Both the Major and the Minor Groove

Juan B. Blanco,[1] M. Eugenio Vázquez,[1] José Martinez-Costas,[2] Luis Castedo,[1] and José L. Mascareñas[1,*]

[1]Departamento de Química Orgánica y Unidad Asociada al CSIC

[2]Departamento de Bioquímica y Biología Molecular Universidad de Santiago de Compostela 15782 Santiago de Compostela Spain

Summary

Attachment of a slightly modified basic region of a bZIP protein (GCN4) to a distamycin-related tripyrrole provides a bivalent system capable of binding with high affinity to specific DNA sequences. Appropriate adjustment of the linker between the two units has led to a hybrid that binds a 9 base-pair-long DNA site (TTTTATGAC) with low nanomolar affinity at 4°C. Circular dichroism and gel retardation studies indicate that the binding occurs by simultaneous insertion of the bZIP basic region into the DNA major groove and the tripyrrole moiety into the minor groove of the flanking sequence. Analysis of hybrids bearing alternative linkers revealed that tight, specific binding is strongly dependent on the length and nature of the connecting unit.

Introduction

The initiation of transcription depends to a large extent on the interaction of proteins called transcription factors with specific DNA sequences located at promoter or enhancer regions of the genes [1, 2]. In many cases, these transcription factors bind to their cognate sequences as noncovalent dimers or multimers, with key DNA contacts responsible for the recognition established by α-helical regions of the protein inserted into the DNA major groove [3–9]. This multimerization strategy ensures tight affinity and high specificity while providing for diversity in target recognition [10–13]. In some cases, such as in several homeodomain-type transcription factors, even the monomers are able to achieve high-affinity DNA binding, although this requires the major groove interaction to be complemented by simultaneous interactions between a separate region of the protein and the minor groove of overlapping or adjacent sequences [14, 15].

In recent years, there have been significant efforts to obtain minimized versions of naturally occurring sequence-specific DNA binding proteins that retain their DNA binding ability, i.e., peptides, which, while being smaller than the natural proteins, are still capable of tight recognition of specific DNA sequences [16]. Although progress in this area has been slow, several designs based on the binding mode of basic-leucine zipper (bZIP) transcription factors have shown interesting DNA binding properties. bZIP proteins bind DNA as leucine zipper-mediated homo- or heterodimers, with the N-terminal basic region (BR) of each monomer inserting into adjacent DNA major grooves [17, 18]. This basic region is largely unstructured in the absence of DNA, but it folds to give an α helix upon specific DNA binding [19, 20]. Although the leucine zipper region does not contact the DNA directly, DNA binding is strongly dependent on such dimerization, as prevention of this process through specific mutations precludes DNA recognition [21, 22]. As one might anticipate from the above comments, monovalent bZIP basic regions exhibit low DNA affinities; however, it has been demonstrated that artificial dimerization of these basic regions, by means of noncovalent or covalent crosslinks, provides peptides that are capable of reproducing the recognition properties of the natural proteins [23–29]. In a recent study, Schepartz et al. showed that grafting the DNA binding residues of the basic region of GCN4 into the preorganized helix of an *aPP* protein leads to miniature proteins of approximately 42 amino acids that are capable of tight monovalent recognition of 5 base-pair (bp)-long specific DNA sites [30, 31].

Results published by Verdine et al. in 1995 showed that 24-mer peptides containing residues of the basic region of GCN4 can fold and bind the target DNA sequence when covalently attached to an appropriate DNA region [32]. Bearing these results in mind, and inspired by the mode of DNA recognition of monomeric homeodomain proteins, we envisaged that suitable tethering of a bZIP-basic region domain to a molecule that binds with moderate to good affinity in the minor groove of a flanking sequence might allow specific delivery of the peptide to the consensus major groove site. More importantly, the resulting conjugate, which would grasp DNA through both the major and the minor groove, should exhibit higher affinity for its target composite DNA site than either of its individual components. Indeed, we have recently found that appropriate crosslinking of a slightly modified basic region of GCN4 to a tripyrrole related to distamycin A provides a bivalent system capable of binding to a designated composite 9 bp DNA site with remarkable affinity [33]. Unfortunately, electrophoretic mobility shift assays of this synthetic conjugate revealed that in addition to the band corresponding to the specific major-minor groove interaction, there is a minor proportion of a slower migrating band. The presence of this less intense band was explained in terms of a competing binding mode in which the oligopyrrole is inserted into the groove but the basic region is disordered and makes nonspecific contacts with the DNA, most probably with the phosphate groups.

Herein, we demonstrate that a slight increase in the length of the tether that links the minor and the major groove binding modules provides for a better affinity, and, more importantly, suppresses the secondary, less specific interaction. The resulting peptide consists of

*Correspondence: qojoselm@usc.es

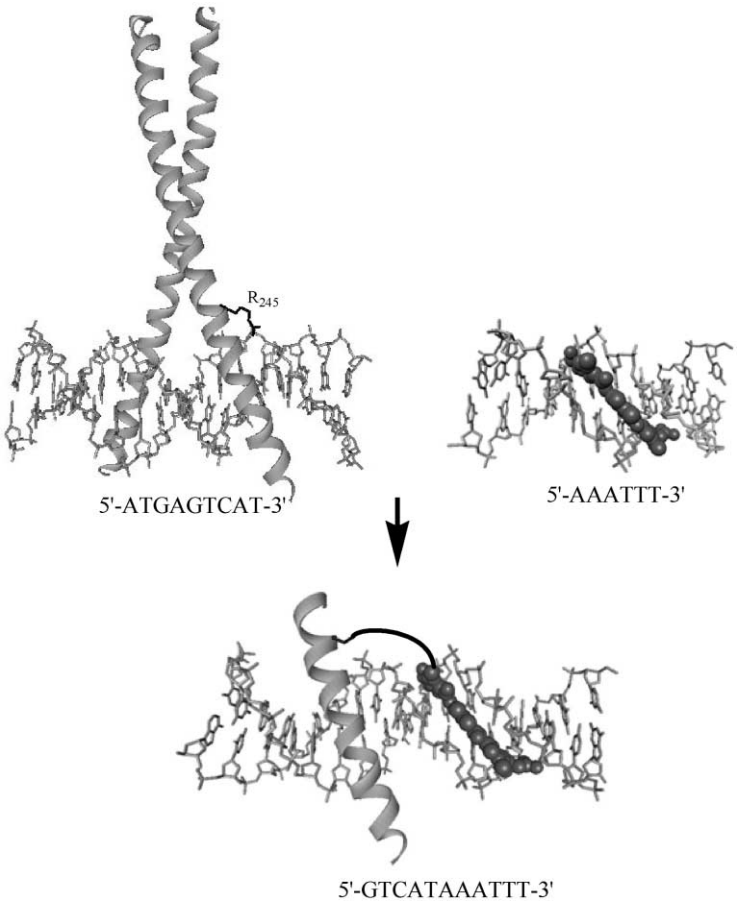

Figure 1. Outline of the Design Strategy

The upper part shows the X-ray structures of GCN4 and distamycin bound to their DNA cognate sites. Arg245, which is mutated to a glutamic acid for coupling to the minor-groove unit, is also indicated. The lower part represents a hypothetical DNA binding model involving the designed peptide-tripyrrole conjugates.

only 22 natural amino acids together with the tripyrrole, and the linker and is, to the best of our knowledge, the highest affinity, smallest peptide derivative described to date that binds specific DNA sequences through the major groove. In addition, it is the only system of all the designed DNA binding peptides that takes part in specific binding in both grooves.

Results

Design and Synthetic Strategy

On the basis of the X-ray structures of the DNA complexes of GCN4 [34] and distamycin A [35] bound to their cognate sequences, we built a hypothetical model for the simultaneous interaction of both molecules on the same face of adjacent DNA sites (Figure 1). The process consisted of superposition of both structures by phosphate alignment until an appropriate arrangement was found to connect the modules. Removal of one of the bZIP chains and the leucine-rich region of the other provided an appropriate three-dimensional model for the design of the linking strategy. Our initial efforts focused on connecting the C-terminal bZIP basic region and the N terminus of the tripyrrole via a peptidic tether, as the syntheses of the corresponding hybrids could be readily carried out in solid phase [36]. Unfortunately, preliminary spectroscopic screenings showed that the resulting peptides failed to bind a variety of potential DNA targets in the desired mode. We realized that this approach might be hampered because the linking tether is forced to depart from the floor of the minor groove. We therefore began to investigate an alternative strategy based on attaching the tether to the nitrogen of one of the distamycin pyrroles, with the aim of exploiting the advantage of the intrinsic preorientation of this unit outwards from the groove. The nitrogen atom of the N-terminal pyrrole and the side chains of amino acids 241 or 245 seemed to provide the most appropriate linkage sites. Since Arg241 is most likely involved in an intramolecular contact with Glu237, which might be relevant for α helix formation, we chose position 245 for connection. It was decided to replace the naturally occurring arginine at that position with a glutamic acid, as it was thought that the latter would offer easier coupling properties.

We also envisioned that the presence of a secondary nitrogen atom in the linking chain might facilitate the phosphate backbone crossover, so our initial designs concerned hybrids **3a**, which contains a relatively tight but apparently long enough linker to span the required crossing distance, and **3b**, which features a slightly longer linker (Figure 2). We also changed the natural amidinium tail of distamycin to a more stable and synthetically accessible dimethylamino group, as it has previously been demonstrated that this change does not significantly alter the binding properties of the molecule [37].

Hybrids **3a** and **3b** were synthesized by coupling the

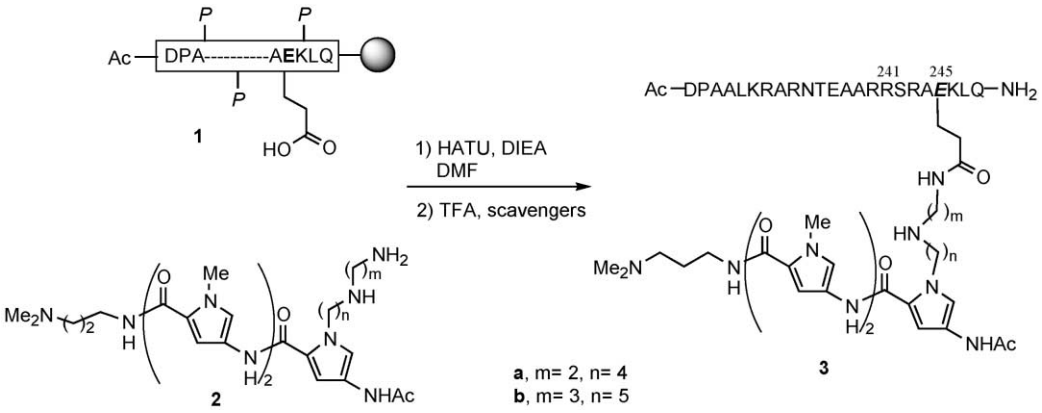

Figure 2. Synthesis of Hybrid Peptides **3a** and **3b**

The gray sphere represents a solid resin. HATU, O-(7-azabenzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate. We use the same amino acid numbering in the peptide as in the natural protein.

side-chain-equipped tripyrroles **2** with the peptide, while this latter is still attached to the resin and fully protected except for the Glu245 residue (Figure 2) [33].

Specific binding of bZIP proteins to DNA is coupled to the folding of the basic region to an α helix. Consequently, circular dichroism (CD) is a particularly useful technique to detect specific interactions, provided that the basic region is not already highly helical in the absence of DNA. Clearly, evidence for an increase in helicity by CD does not necessarily correlate with high affinity binding, as the concentrations required for the CD experiments are in the micromolar range; however, tight, specific binding does require the induction of a considerable proportion of an α-helical structure, a characteristic that can be readily detected by observation of the change in the negative ellipticity at 222 nm in the CD spectrum.

As shown in our previous communication [33], addition of an 18 bp duplex containing the designated hybrid DNA sequence (T/CRE^{hs}) (the target site of this sequence does not match exactly with that of the model, but it was used on the basis of the preferred (1:1) recognition sites for distamycin: [38]) to **3a** at 4°C did not produce a significant variation of the CD signal at 222 nm; however, we did observe a considerable positive signal at 330 nm. These results were explained in terms of a peptide-DNA interaction in which the tripyrrole is inserted into the minor groove [39], but the basic region has not accomplished the folding binding process characteristic of specific binding. Such a failure could arise from the tether being too short to allow an appropriate strain-free simultaneous docking of the tripyrrole and the peptide into the grooves. Indeed, addition of the target duplex DNA T/CRE^{hs} to **3b**, which features a longer linker, induced a substantial increase in the magnitude of the negative signal at 222 nm. This result suggests that in this case the basic peptidic domain is folding and inserting into the DNA major groove (Figure 3A). The band at 330 nm, which should arise from minor-groove binding of the oligopyrrole, was also clearly observed. Significant increases in the ellipticity were not observed when **3b** was incubated with TcgCRE^{hs}, a dsDNA that contains both binding sites separated by two base pairs, nor with a duplex DNA featuring a base pair mismatch at the basic region binding subsite (T/CRE^{hs}*m*). It should also be noted that a significant proportion of the increase in the helicity of **3b** upon addition of the target DNA is lost when the complexation is carried out at room temperature.

In order to confirm that the increase in negative ellipticity at 222 nm observed for hybrid **3b** upon incubation with the cognate duplex oligonucleotide correlates with specific binding and strong affinity, its DNA binding properties were studied by gel mobility shift analysis. These binding studies were carried out in 20 mM Tris-HCl (pH 7.5), 100 mM KCl, 2 mM $MgCl_2$, and 2 mM EDTA, all at 4°C. As illustrated in Figure 3B, nanomolar concentrations of peptide **3b** produced mobility-retarded bands corresponding to the expected 1:1 complexes with the duplex T/CRE^{hs}. However, two bands are observed, the faster migrating one being the most intense. Incubation of peptide **3b** with a dsDNA mutated at the BR binding site (T/CRE^{hs}*m*) gave a band with similar mobility to the slower migrating band in the preceding experiment (lanes 7–11, Figure 3B). Incubation of the target DNA T/CRE^{hs} with hybrid **3a**, which did not exhibit the α-helical folding transition, gave a band of similar mobility to the slower migrating band obtained with **3b** (lanes 15–17, Figure 3B).

These results, together with the CD data, can be explained by assuming that the more intense, faster migrating band obtained upon incubation of **3b** with the target composite dsDNA corresponds to the expected compact major-minor binding mode. On the other hand, the less intense band results from a complex in which the oligopyrrole is inserted into the groove, but the basic peptide makes only nonspecific contacts with the DNA (most probably and mainly with the phosphates). The high affinity of this less specific binding is consistent with previous studies by Bruice and coworkers concerning microgonotropens—tripyrrole derivatives bearing positively charged N-alkylamine chains that show surprisingly high DNA affinities [40].

Synthesis and DNA Binding Properties of Hybrid 5

The dramatic change observed in the DNA binding ability of the above hybrid peptides by slightly enlarging the linker highlights the importance of having sufficiently

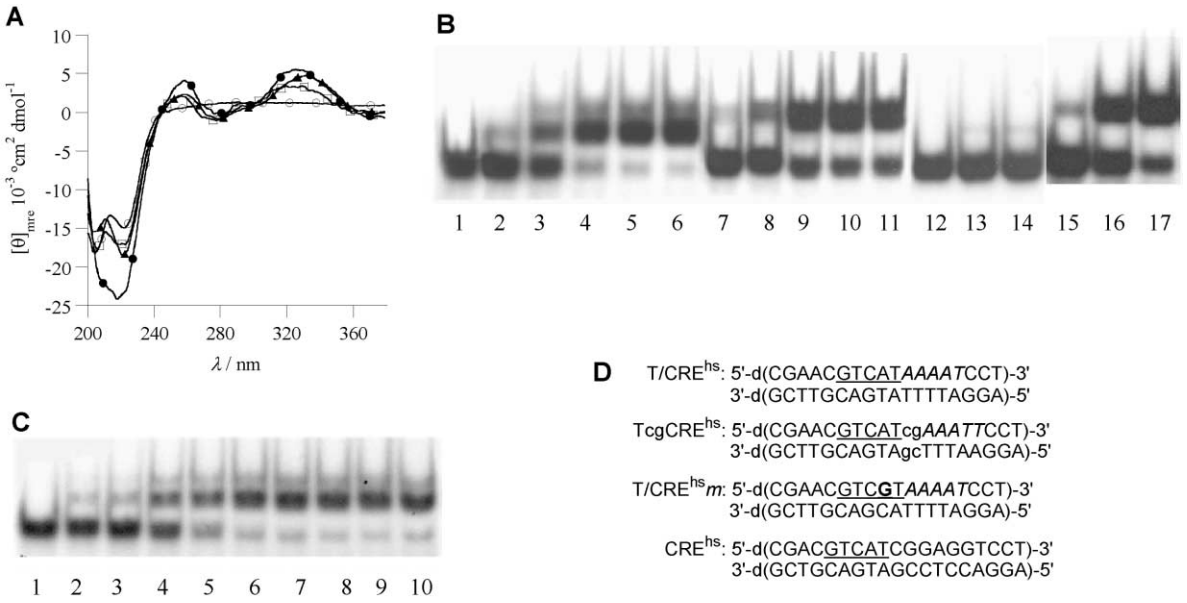

Figure 3. DNA Binding Properties of **3b**

(A) CD difference spectra of peptide **3b** in the presence or absence of ds-oligonucleotides: in the absence of DNA (○), in the presence T/CRE^{hs} (●), in the presence TcgCRE^{hs} (▲), in the presence of T/CRE^{hs}*m* (□). CD spectra were obtained at 4°C as described in Experimental Procedures and were slightly smoothed to facilitate viewing. The contribution of the DNA has been subtracted.

(B and C) Autoradiograms showing the binding of hybrids **3** to ^{32}P DNAs: (B) using ~1 nM ^{32}P DNA, lanes 1–6: T/CRE^{hs}, [**3b**]: 0, 7.7, 19, 38, 58, 77 nM; lanes 7–11: T/CRE^{hs}*m*, [**3b**]: 7.7, 19, 38, 58, 77 nM; lanes 12–14: CRE^{hs}, [**3b**]: 38, 77, 154 nM; lanes 15–17: T/CRE^{hs}, [**3a**]: 7.7, 38, 77 nM. (C) Using ~45 pM ^{32}P DNA, lanes 1–10: T/CRE^{hs}, [**3b**]: 0, 1, 2, 5, 10, 20, 40, 60, 80, 100 nM.

(D) Sequences of duplex oligonucleotides used. The BR subsite (CRE^{hs}) is underlined, and the tripyrrole subsite (T) is in italics.

long linkers that avoid the introduction of strain when both partners fit into their cognate sites. However, long and flexible linkers can contribute unfavorably to the entropic cost of the process. Therefore, in trying to optimize our design, we decided to slightly elongate the linker by adding a glycine unit, which represents an increase in length of ~2.5–3 Å. The synthesis of this new hybrid was readily achieved using the route outlined in Figure 4. Treatment of a DMF solution of tripyrroleamine **2b** with Boc-Gly in the presence of HBTU and DIEA followed by removal of the Boc group in an acidic medium provided amine **4** in 50% yield. This compound was coupled to the peptide through a similar procedure to that used in the case of the aforementioned amines **2**. This route led to the expected hybrid **5** (26% isolated yield, taking also into account the peptide synthesis).

As can be deduced from Figure 5A, the CD spectrum of the peptide derivative **5** in the presence of the target DNA revealed a slightly higher helicity than that observed for its relative **3b**. The expected positive band at 330 nm that corresponds to the tripyrrole chromophore was also observed. The CD spectra acquired in the presence of a duplex oligo having mutated BR half-sites (T/CRE^{hs}*m*) or binding sites separated by 2 bp (TcgCRE^{hs}) do not show a significant α-helical folding transition.

The CD spectrum of the DNA in the complex **5**-T/CRE^{hs} still shows typical bands of a B-DNA form (minima at 247 nm and 210 nm and a maximum at 275 nm; see Supplemental Data at http://www.chembiol.com/cgi/content/full/10/8/713/DC1), which suggests that protein binding does not promote significant changes in the

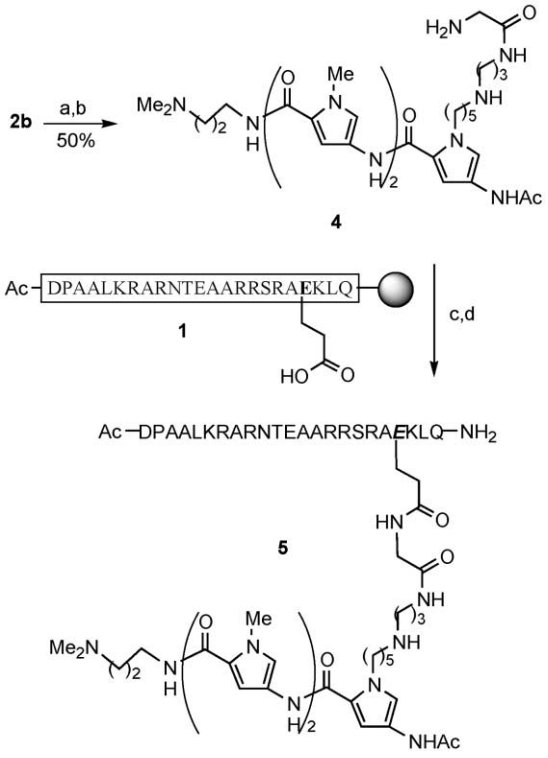

Figure 4. Synthesis of Peptide-Tripyrrole Conjugate **5**

a, Boc-Gly, HBTU, DIEA/DMF; b, TFA, CH_2Cl_2; c, HATU, DIEA/DMF, **1** (1.1 equiv.); d, TFA, EDT, PhOH, thioanisole, H_2O. HBTU, 2-(1*H*-benzotriazol-1-yl)-1,1,3,3,-tetramethyluroniumhexafluorophosphate.

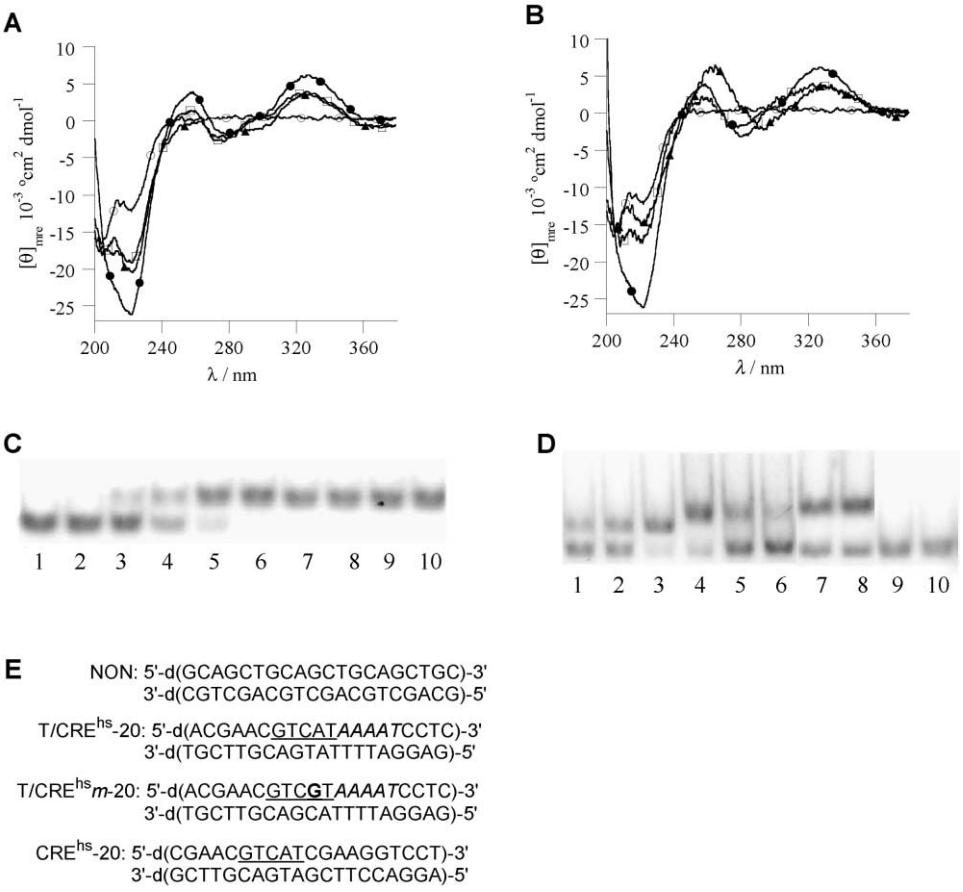

Figure 5. DNA Binding Properties of Hybrid **5**

(A) CD difference spectra of peptide **5** in the absence of DNA (○), in the presence T/CRE^{hs} (●), in the presence TcgCRE^{hs} (▲), and in the presence of T/CRE^{hs}*m* (□).

(B) CD difference spectra of peptide **5** in the absence of DNA (○), in the presence T/CRE^{hs} (●), in the presence NON (▲), and in the presence of CRE^{hs}-20 (□).

(C) Lanes 1–10: T/CRE^{hs}, [**5**]: 0, 1, 2, 5, 10, 20, 40, 60, 80, 100 nM.

(D) Mobility shift comparison of complexes of **5** with different DNAs: lanes 1–3, T/CRE^{hs}-20 [**5**]: 2, 5, 10 nM; lanes 4–6, T/CRE^{hs}*m*-20 [**5**]: 100, 10, 5 nM; lanes 7 and 8, TcgCRE^{hs}, [**5**]: 40, 80 nM; lanes 9 and 10, CRE^{hs}-20 [**5**]: 80,160 nM.

(E) Sequences of other duplex oligonucleotides used in the above experiments.

DNA conformation. As shown in Figure 5B, addition of a random, nonspecific DNA (NON) or a duplex oligo containing only the BR subsite (CRE^{hs}-20) to **5** does not appreciably alter the ellipticity at 220 nm. The presence of a positive signal at 330 nm suggests that at the concentrations used for the CD experiment there is a partial formation of a complex in which the tripyrrole moiety is bound to the duplex.

It was gratifying to find that titration gel shift experiments on T/CRE^{hs} with increasing amounts of **5** revealed a single binding mode in which half-maximal binding occurs at concentrations as low as 5 nM at 4°C (Figure 5C). As expected, the use of DNAs bearing mutated peptide binding half-sites [T/CRE^{hs}*m*-20] or bearing a 2 bp spacing between the binding subsites (TcgCRE^{hs}) gave rise only to the less specific slower migrating complex (Figure 5D, lanes 4–6 and 7–8, respectively. Note that in this experiment slightly longer oligos were used in order to observe more clearly the shift differences among bands). EMSA experiments also showed that the affinity of **5** for the nonspecific random DNA NON is lower than 1 μM (data not shown) and that the binding of **5** to T/CRE^{hs} is not very much affected by the presence of excess of this scrambled dsDNA (see Supplemental Data at above URL). On the other hand, the gel shift assay revealed that the interaction of **5** with a duplex DNA containing just the BR binding subsite (CRE^{hs}-20) is very weak (affinity lower than 200 nM; lanes 9 and 10, Figure 5D).

In the absence of the nonspecific DNA competitor, we have calculated an apparent dissociation constant of $3 \pm 0.2 \times 10^{-9}$ M at 4°C for the complex **5**-T/CRE^{hs} (Figure 6), whereas the affinity of **5** for a dsDNA bearing a mutated bp at the BR binding site (T/CRE^{hs}*m*) is around 15 nM (It is noteworthy that in this case full binding saturation is not reached even at 120 nM; see Supplemental Data at above URL). The affinity of the hybrid for its composite designated site is considerably higher

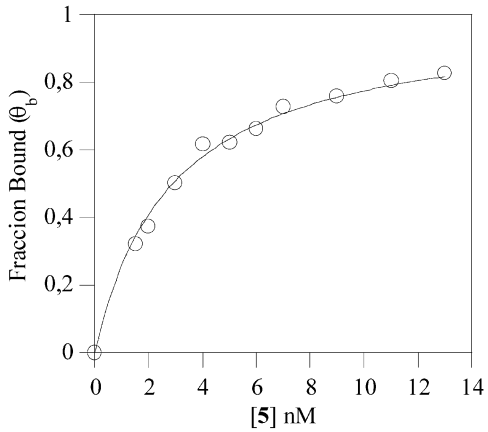

Figure 6. PAGE Titration of Hybrid **5**

A graphical representation of the data with nonlinear least-squares fit to the equation $\theta_b = P/(K_d + P)$, where P is the concentration of **5** and K_d is the dissociation constant.

than that of any of its components for their respective sequences, as the basic region peptide alone binds with low micromolar affinities [30], and the tripyrrole **4** shows an apparent dissociation constant of ~0.5 μM as deduced from a CD equilibrium binding isothermal curve (see Supplemental Data at above URL).

Synthesis and DNA Binding Properties of a Hybrid with a Linker Containing a Triglycine Unit

The above results suggest that the linker in hybrid **5** has an appropriate length to allow the required bivalent DNA binding. Given this information, it was intriguing to find out if other linkers that span the same distance might also lead to similar outcomes. We therefore examined the DNA binding properties of the tripyrrole-peptide conjugate **8**, a derivative with a linker of similar length to that in compound **5** but lacking the middle secondary nitrogen and bearing a relatively rigid triglycine unit.

The synthesis of the required tripyrrole system was achieved by alkylation of **6** with Boc-3-iodopropylamine, removal of the Boc, and coupling of the resulting amine to commercially available Z-triglycine (48% yield; Figure 7A). This amine was attached to the basic region peptide using the same conditions as before to give the expected conjugate **8** (~34% yield). Remarkably, the CD spectrum of **8** in the presence of the cognate duplex DNA T/CRE^{hs} showed a weak increase in helicity, much weaker than that observed with the parent hybrid **5** (Figure 7B). PAGE experiments confirmed that this compound is unable to specifically bind to the designated sequence, and, curiously, the formation of the DNA-peptide complexes cannot be observed until reaching peptide concentrations of 60 nM (Figure 7C).

Discussion

There is a great deal of interest in obtaining artificial mimics of transcription factors that although smaller in size are yet capable of reproducing their DNA binding specificity and affinity. It is desirable that these artificial DNA binding peptides are able to address sequences containing a relatively high number of bases so that they could target particular sites in the genome with higher specificity. These molecules, in addition to providing basic information about the factors that govern sequence-specific protein-DNA interactions, might find important biomedical applications as designed genome interference agents [41].

Although progress in this area has been slow and intermittent, the last decade has seen remarkable achievements, particularly in the field of zinc finger and bZIP transcription factors. This latter type of transcription factor recognizes DNA through a bipartite structural motif that consists of a coiled-coil leucine zipper dimerization domain and a positively charged basic region that directly contacts DNA in the major groove by folding into an α helix. Monovalent bZIP basic regions exhibit very low sequence-specific DNA affinities unless they are engineered to be appropriately preorganized in an α-helical form [30, 31].

Inspired by naturally occurring homedomain DNA binding proteins that are capable of recognizing specific DNA sequences in a monomeric form and contact DNA through both the major and minor groove, we studied the DNA binding potential of covalent conjugates between major-groove binding bZIP basic regions of GCN4 and tripyrrole minor-groove recognition units related to distamycin A.

Early DNA binding assays carried out with hybrids resulting from a linear peptidic connection between the N-terminal amine of the tripyrrole and the C terminus of the basic region of GCN4 [36] revealed that the peptide is unable to fold to the α helix required for specific binding. Inspection of qualitative molecular models obtained from the crystallographic structures of the DNA complexes of GCN4 and distamycin A suggested the pyrrolic nitrogen of the tripyrrole as a more appropriate tethering position, as it is already projecting outward from the minor groove.

The CD spectra of our initial synthetic construct (**3a**) in the presence of a duplex oligonucleotide containing the designated binding sequence were consistent with insertion of the tripyrrole unit into the DNA minor groove, as deduced from the appearance of a new positive band at 330 nm. However, the absence of an increase in negative ellipticity at 222 nm is indicative of the basic region failing to fold and insert into the major groove. We reasoned that this failure might be due to geometrical restrictions imposed by a linker that is possibly too short to ensure that the binding units can appropriately reach their respective sites. We therefore investigated whether a small increase in the length could provide better results. Indeed, addition of the designated composite dsDNA to the hybrid **3b**, which contains a couple of additional methylene groups in the linker, resulted in a large increase in the negative ellipticity at 222 nm, with θ_{mrw} changing from −10,000 to −24,000 deg.cm^2/dmol, a change consistent with the formation of a considerable α-helical structure at 4°C (from ~40% to 84% of helicity increase). The spectrum also provides evidence of a strong positive band induced at 330 nm that must arise from insertion of the tripyrrole unit into the minor groove. Control CD experiments in the presence of duplex oligonucleotides containing a different half-site spacing

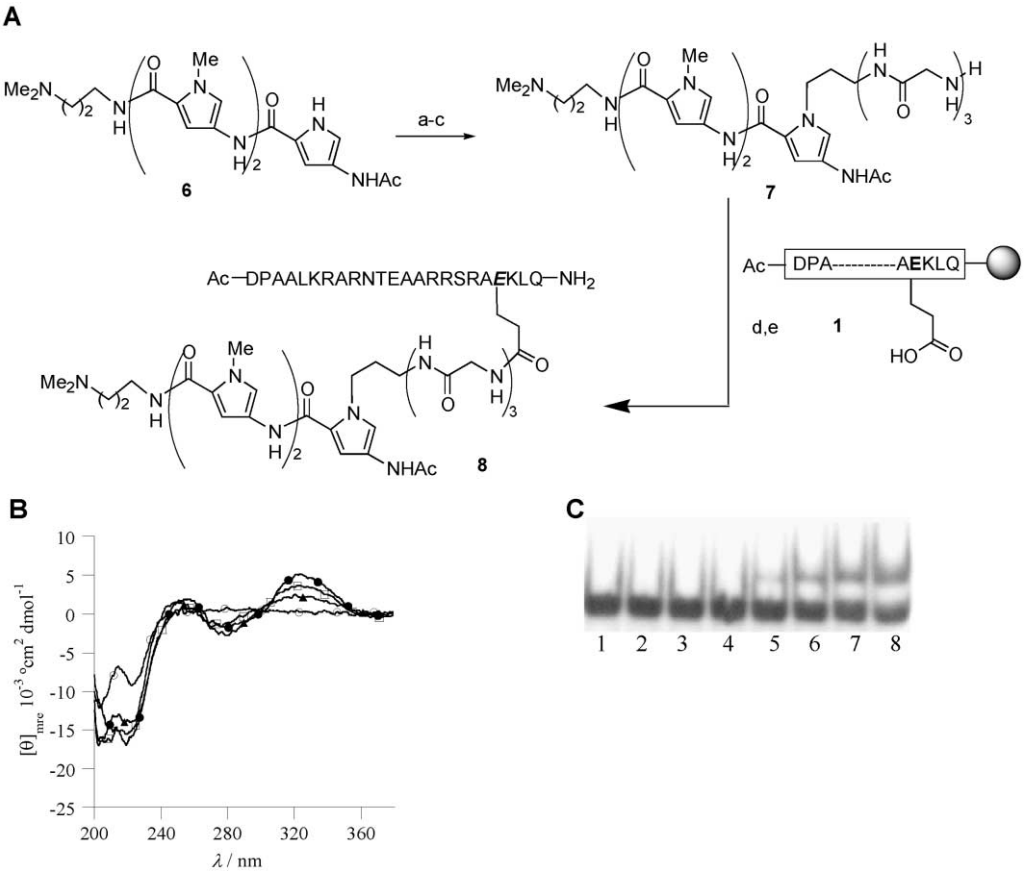

Figure 7. Synthesis and DNA Binding Properties of Peptide-Tripyrrole Conjugate 8

(A) a, I-$(CH_2)_3$-NHBoc, K_2CO_3, refluxing acetone; b, i, TFA, CH_2Cl_2; ii, Z-$(Gly)_3$, HBTU, DIEA/DMF; c, H_2, Pd-C, MeOH; d, HATU, DIEA/DMF, **1**; e, TFA, EDT, PhOH, thioanisole, H_2O.

(B) CD difference spectra of peptide **8** in the presence or absence of ds-oligonucleotides. CD in the absence of DNA (○), in the presence T/CRE^{hs} (●), in the presence TcgCRE^{hs} (▲), in the presence of T/CRE^{hs}*m* (□).

(C) Autoradiogram showing the binding of hybrid **8** to ^{32}P DNAs (~45 pM). Lanes 1–8: T/CRE^{hs}, [**8**]: 0, 5, 10, 20, 40, 60, 80, 100 nM.

(TcgCRE^{hs}) or a mutated peptide binding site (T/CRE^{hs}*m*) revealed the formation of a significantly lower proportion of α helix, which suggests that specific binding requires that both cognate half-sites are present in an appropriate flanking arrangement. Polyacrylamide gel shift retardation experiments confirmed that **3b** binds the designated hybrid site with nanomolar affinity. However, in addition to the bandshift that must correspond to the desired compact, specific complex, there is a less intense band of lower mobility. This latter band has a similar mobility to the band obtained when the peptide was incubated with a dsDNA mutated at the peptide binding site (T/CRE^{hs}*m*). This result, together with the dichroism data, suggests that this band corresponds to a complex in which the tripyrrole module is specifically bound, and the basic peptidic domain is most probably involved in nonspecific electrostatic contacts with the DNA backbone. The relatively high affinity of this less specific binding mode is not surprising bearing in mind previous results obtained with microgonotropenes, tripyrroles that are equipped with N-alkylamine side chains and are protonated at neutral pH [40], and considering that GCN4 can bind with relatively high affinity to DNA sites containing only a single consensus half-site [42].

The remarkable improvement in DNA affinity achieved after slight elongation of the linker (from **3a** to **3b**) raised the question of whether a further increase in length might lead to a higher affinity and hence suppress the formation of the competing, less specific interaction. The search for optimal linkers that maximize polyvalent interactions is not trivial because of the partial compensating effects of entropy and enthalpy: conformational flexibility favors the likehood that the interaction occurs without strain but increases the conformational entropic cost of the association [43]. We therefore decided to assess the consequences of lengthening the linker by adding a single glycine unit.

We were pleased to find that the new hybrid **5** exhibited a slightly tighter binding than **3b** to the consensus composite site T/CRE^{hs} and, more importantly, in this case, the competing less specific binding mode is not detected at all. As expected, the use of a DNA mutated at the peptide binding site (T/CRE^{hs}*m*) leads exclusively to the less specific complex. CD spectroscopy confirmed that the high-affinity interaction with T/CRE^{hs} conveys the expected folding transition of the basic region to an α helix. An electrophoresis mobility shift titration gave a value for the dissociation constant of ~3 nM at

4°C for the interaction of **5** with T/CRE^{hs}. This result represents at least a 150-fold increase in binding affinity relative to the unlinked units.

The CD spectrum of **5** in the presence of a dsDNA lacking the consensus tripyrrole binding half-site (CRE^{hs}-20) or one containing a random sequence (NON) showed no significant degree of helical transitions. Albeit the presence of a weak positive signal at 330 nm could indicate a certain degree of affinity, EMSA experiments confirmed that **5** binds to both DNAs with very low affinity, as complexes were not detected even at concentrations of 160 nM. In terms of specificity, it is clear that **5** binds to the designated match site more than 1000-fold better than to the other sites that have a noncognate sequence for the tripyrrole (CRE^{hs} and NON) and at least five times better than to regions bearing a consensus distamycin A-T-rich half binding site but a mutated base pair at the peptide binding sequence (T/CRE^{hs}*m*).

It has been proposed that peptidic-type tethers can be particularly useful for linking bivalent systems as a consequence of a minor conformational entropy cost of the linker upon interaction between the systems and their receptors [43, 44]. We therefore considered it of interest to examine the DNA binding properties of hybrid **8**, which contains a propyltriglycine tether spanning approximately the same length as the pentylpropylglycine present in **5**. Quite surprisingly, this hybrid showed a much lower affinity to T/CRE^{hs} than **5**, even when compared with the aforementioned half-specific binding of hybrids **3** and **5**. Although the nature of the tether in **8** is quite different to that in **5**, the above results seem to point to the secondary nitrogen present in the linker of **3b** and **5** as a very important structural element for obtaining successful DNA binding peptides using our minor-major groove binding approach. In any case, further modifications of the linker are needed to establish the ultimate factors that determine the DNA-recognition properties of these types of synthetic DNA binders.

Significance

Designing miniature proteins that can mimic the sequence-specific DNA recognition of DNA binding proteins is a major goal of current research at the chemistry-biology interface. In particular, there is great interest in obtaining molecules that can specifically address long DNA sequences as potential selective genome tools in transcription therapy. Although an extensive number of DNA binding proteins recognize DNA by means of tandem segments that interact simultaneously with both the major and the minor grooves of the target DNA duplex, this binding mode has not been investigated in artificially designed DNA binding molecules. Herein, we have demonstrated that complementing the DNA binding ability of α-helical basic domains of naturally occurring transcription factors with minor groove binding units capable of interacting with moderate to good affinity in flanking DNA sequences provides a novel and efficient DNA-recognition strategy. By appropriately adjusting the segment that links the major and the minor groove binding modules, we have obtained a synthetic derivative that binds to relatively long DNA sites (9–10 bp) with low nanomolar affinity, thus representing the smallest transcription factor-based peptide derivative designed hitherto capable of recognizing specific DNA sequences with such high affinity. We have also shown that the linker is not a mere spectator but is actively engaged in the recognition process, since changing its structure while keeping the length drastically reduces the DNA affinity of the system.

Taking into account that this new type of bipartite minor-major groove binding can raise interesting interference effects with DNA information processing at an early stage of gene expression, we predict that this new approach for DNA recognition could find important biological and medical applications. Work is under way to obtain structural details of the peptide-DNA complex and to improve the design process so that high affinity and specific binding can be reached at higher temperature.

Experimental Procedures

Peptide Synthesis

Peptide synthesis was performed using standard Fmoc-solid phase synthesis method on a Rink-MBHA amide resin (~0.46 mmol/g), using mixtures of HBTU/HOBt as coupling agents, DIEA as base, and DMF as solvent. The cleavage/deprotection step was performed by treatment of the resin-bound peptide with the following mixture: 830 μl TFA, 25 μl EDT, 50 mg PhOH, 50 μl tioanisole, and 50 μl H_2O (300 μl of this mixture for each 10 mg of resin). The amino acids used in the synthesis were standard protected, except the Fmoc(All)-Glu-OH introduced at position 245. The cleaved peptides were analyzed and purified by RP-HPLC with a LiChrospher WP 300 RP-18 250 × 4 column for analytical experiments and Nucleosil 120-10 C-18 250 × 8 for semipreparative purifications and a linear gradient of CH_3CN in H_2O containing 0.1% TFA (v/v).

Circular Dichroism Assays

CD measurements were made at 4°C (unless otherwise stated) in 2 mm cell. Samples contained 10 mM phosphate buffer (pH 7.5), 100 mM NaCl, 5 μM peptide, and 5 μM ds-oligo when present. The peptide-DNA mixtures were incubated for 15 min at 4°C before registering. The spectra are the average of ten scans and were slightly smoothed using the "smooth" macro implemented in the program Kaleidagraph (v 3.5, Synergy Software). Spectra of the peptides in the presence of DNA were calculated as the difference between the spectra of the peptide-DNA mixture and the spectrum of free DNA, unless otherwise stated.

PAGE Experiments

For gel mobility shift assays, binding reactions were performed over 30 min at 4°C using ~45 pM ^{32}P-labeled dsDNAs (unless otherwise stated) in a binding mixture (20 μl) containing 20 mM Tris (pH 7.5), 100 mM KCl, 2 mM $MgCl_2$, 2 mM EDTA, 10% glycerol, 0.3 mg/ml BSA, and 2% NP-40. Products were resolved by PAGE using a 10% nondenaturing acrylamide gel and 0.5× TBE buffer and analyzed by autoradiography (0.5× TBE buffer is 32 mM boric acid, 50 mM Tris base, and 1 mM EDTA [pH 8]).

Synthesis of Tripyrroles 4 and 7

To a solution of tripyrrole **2b** (80 mg, 0.125 mmol) and DIEA (36.8 μl, 0.215 mmol) in DMF (1 ml) was added a freshly prepared mixture of Boc-gly-OH (22 mg, 0.125 mmol) and HBTU (47 mg, 0.125 mmol). The resulting solution was stirred for 1 hr at room temperature, and the solvents were evaporated under vacuum. The resulting residue was purified by semipreparative RP-HPLC (gradient 5 to 95%, 30 min; retention time [R_t] = 12.76 min). The coupled tripyrrole-glycine product was dissolved in CH_2Cl_2 (1 ml) and cooled to 0°C. TFA (1 ml) was added dropwise, and the resulting orange solution was stirred at 0°C for 1 hr and at room temperature for another 2 hr.

Solvents were removed under vacuum at room temperature, and residual TFA was removed by codistillation with CH_2Cl_2. The solid residue was identified as the desired product **4** (44 mg, 50%, mp = 130°C), (analytical RP-HPLC, gradient 5%–95%, 30 min, R_t = 11.4 min). ^{1}H NMR δ (CD_3OD): 1.26–1.29 (m, 2H), 1.35–1.42 (m, 2H), 1.64–1.73 (m, 2H), 1.80–1.90 (m, 2H), 1.96–2.02 (m, 2H), 2.08 (s, 3H), 2.91 (s, 6H), 2.94–3.03 (m, 4H), 3.17 (t, *J* = 7.3 Hz, 2H), 3.33–3.43 (m, 4H), 3.68 (s, 2H), 3.89 (s, 3H), 3.91 (s, 3H), 4.34 (t, *J* = 7.2 Hz, 2H), 6.80 (s, 1H), 6.89 (s, 1H), 6.90 (s, 1H), 7.18 (s, 1H), 7.21 (s, 1H), 7.24 (s, 1H). ^{13}C NMR δ (CD_3OD): 22.9 (CH_3), 24.4 (CH_2), 26.6 (CH_2), 26.7 (CH_2), 27.4 (CH_2), 32.1 (CH_2), 36.8 (CH_3), 37.0 (CH_2), 37.1 (CH_2), 42.6 (CH_2), 43.7(CH_3), 46.2 (CH_2), 48.8 (CH_2), 49.1 (CH_2), 49.8 (CH_3), 56.8 (CH_2), 106.3 (CH), 106.4 (CH), 106.6 (CH), 115.8 (C), 119.7 (CH), 120.7 (CH), 120.9 (CH), 123.2 (C), 123.3 (C), 123.8 (C), 124.1 (C), 124.5 (C), 161.2 (C), 161.3 (C), 164.6 (C), 170.4 (C), 170.9 (C). HRMS: calculated for $C_{34}H_{53}N_{11}O_5$ 695.42311, found 695.42262. UV(H_2O), λ_{max} (ε) 304 (32274).

To a solution of the trypyrrole **6** (100 mg, 0.202 mmol) in dry acetone (5 ml) was added dry K_2CO_3 (300 mg) and 1-iodo-*tert*-butyl-2-aminopropylcarbamate (288 mg, 1.01 mmol). The reaction mixture was refluxed for 5 hr. The suspension was filtered through celite, and the filtrate was concentrated. The product was purified by flash chromatography (aluminium oxide, 5% MeOH, CH_2Cl_2). The alkylated compound was dissolved in CH_2Cl_2 (3 ml) and cooled to 0°C. TFA (3 ml, 15 min) was added dropwise, and the resulting orange solution was stirred at 0°C for 1 hr and at room temperature for another 2 hr. Solvents were removed in vacuo at room temperature, and residual TFA was removed by codistillation with CH_2Cl_2. The residue was identified as the desired amine (90 mg, 80%). *N2*-(5-{[3-(dimethylamino) propyl]carbamoyl}-1-methyl-1*H*-3-pyrrolyl)-4-{[1-(3-aminopropyl)-4-methylcarboxamido-1*H*-2-pyrrolyl]carboxamido}-1-methyl-1*H*2-pyrrolecarboxamide: ^{1}H NMR δ (CD_3OD): 1.63–1.87 (m, 7H), 2.67–2.75 (m, 8H), 2.98–3.11 (m, 4H), 3.50 (s, 3H), 3.52 (s, 3H), 3.93 (t, *J* = 6.6 Hz, 2H), 6.56 (s, 1H), 6.60 (s, 1H), 6.61 (s, 1H), 6.82 (s, 1H), 6.85 (s, 1H), 6.90 (s, 1H). ^{13}C NMR δ (CD_3OD): 22.0 (CH_2), 22.9 (CH_3), 24.2 (CH_2), 36.8 (CH_3), 37.0 (CH_2), 37.5 (CH_2), 51.4 (CH_3), 62.0 (CH_2), 63.8 (CH_2), 104.3 (CH), 106.5 (CH), 106.6 (CH), 113.7 (C), 118.2 (C), 122.9 (CH), 123.0 (CH), 124.0 (C), 124.5 (C), 125.1 (C), 160.6 (C), 161.3 (C), 164.4 (C), 170.6 (C). HRMS: calculated for $C_{27}H_{39}N_9O_4$ 553.31250, found 553.31160.

To a solution of this amine (90 mg, 0.163 mmol) and DIEA (36.8 μl, 0.215 mmol) in DMF (2 ml) was added a freshly made mixture of Z-triglycine acid (103 mg, 0.320 mmol) and HBTU (121 mg, 0.320 mmol). The resulting solution was stirred for 1 hr, and the solvents were evaporated in vacuo. The resulting residue was purified by semipreparative RP-HPLC (gradient 5%–95%, 30 min, R_t = 16.88 min, 98 mg, 70%). The compound was dissolved in MeOH (20 ml) and hydrogenated for 2 hr over 10% palladium on charcoal (60 mg) at room temperature (balloon pressure). The catalyst was removed by filtration through celite, and the filtrate was concentrated to give the desired product **7** as a yellow solid (78 mg, 60%, mp = 198°C). Analytical RP-HLPC: gradient 5%–95%, 30 min, R_t = 10.73 min. ^{1}H NMR δ (CD_3OD): 1.91–2.04 (m, 4H), 2.08 (s, 3H), 3.05 (s, 6H), 3.31 (m, 6H), 3.54 (s, 2H), 3.82 (s, 2H), 3.86 (s, 3H), 3.89 (s, 3H), 3.93 (s, 2H), 4.11 (s, 2H), 6.87 (s, 1H), 6.93 (s, 1H), 6.98 (s, 1H), 7.17 (s, 1H), 7.19 (s, 1H), 7.22 (s, 1H). ^{13}C NMR δ (CD_3OD): 22.9 (CH_3), 23.8 (CH_2), 24.2 (CH_2), 36.8 (CH_3), 36.9 (CH_2), 43.3 (CH_2), 43.8 (CH_2), 43.9 (CH_2), 49.8 (CH_3), 51.5 (CH_3), 62.9 (CH_2), 63.3 (CH_2), 104.0 (CH), 106.5 (CH), 114.4 (CH), 115.9 (C), 120.6 (CH), 120.7 (CH), 120.8, (CH) 123.2 (C), 123.3 (C), 124.2 (C), 124.6 (C), 125.3 (C), 160.6 (C), 161.3 (C), 162.7 (C), 164.3 (C), 170.4 (C), 172.2 (C), 172.6 (C). HRMS: calculated for $C_{33}H_{48}N_{12}O_7$ 724.37689, found 724.37642. UV(H_2O), λ_{max} (ε) 304 (32274).

Synthesis of the Solid Phase-Bound Peptide 1

After introducing all of the 23 Fmoc amino acids, the resin-bound peptide was treated with a solution of piperidine (5 ml, 20% in DMF) for 45 min. After DMF washings (3 × 5 ml for 5 min), the acetylation was carried out immediately by treating the resin with DIEA (2.5 ml, 0.2 M in DMF) and Ac_2O (5 ml, 20% in DMF) for 1 hr. After filtration, the resin was washed with DMF (3 × 5 ml for 5 min), *i*PrOH (5 ml for 5 min), and Et_2O (5 ml for 5 min). HPLC analysis of the residue obtained after deprotection/cleavage of a resin aliquot showed a majoritary peak (analytical column, gradient 10% to 35%, 30 min, R_t= 17.36 min) which corresponds to the expected acetylated peptide, fully deprotected except of the allyl group of glu245 side chain, as confirmed by ESI-MS [MH^+] m/z calculated for $C_{112}H_{198}N_{43}O_{34}$ 2691.0, observed 2690.6.

The selective deprotection of the allyl group on the acetylated resin-bound peptide was carried out by treating the resin (160 mg) with 2% H_2O/CH_2Cl_2 (6 ml), morpholine (0.42 ml, 4.84 mmol), and $Pd(PPh_3)_4$ (29 mg, 0.025 mmol). The mixture was smoothly shaken under argon for 12 hr, and the resin was filtered and washed with DMF (3 × 5 ml for 5 min), *i*PrOH (2 × 5 ml for 5 min), and CH_2Cl_2 (5 ml for 5 min). The formation of the resin-bound peptide **1** was confirmed by cleavage/deprotection of an aliquot and ESI-MS analysis of the majoritary HPLC peak (analytical column, 10% to 35%, 30 min, R_t = 16.99 min): [MH^+] m/z calculated for $C_{109}H_{194}N_{43}O_{34}$ 2649.4, observed 2650.0.

Synthesis of the Peptide-Tripyrrole Hybrids

The following is the general procedure for the synthesis of **3a**, **3b**, **5**, and **8**, exemplified for the coupling of tripyrrole **2b** to the resin-linked peptide **1** to give hybrid **3b**: DMF (1 ml) was added over 20 mg of **5** (Eppendorf tube), and the suspension was shaken for 1 hr to ensure a good resin swelling. The DMF was removed and 150 μl of a 0.5 M solution of HATU in DMF and DIEA (150 μl, 0.5 M in DMF) was added. The resulting mixture was shaken for 5 min, and 100 μl of a 0.5 M solution of the tripyrrole **2b** (10 mg in 100 μl) and DIEA (100 μl, 0.5 M in DMF) was added. The reaction mixture was shaken for 2 hr, and the resin was washed with DMF (3 × 0.6 ml for 5 min), *i*PrOH (1 × 0.6 ml for 5 min), and finally Et_2O. Cleavage/deprotection with standard conditions afforded a major product that was purified by RP-HPLC (gradient 10% to 35%, R_t= 21.91 min). ESI-MS analysis confirmed the formation of the desired hybrid **3b** (~33% yield, considering also the peptide synthesis). [MH^+] m/z calculated for $C_{141}H_{242}N_{53}O_{37}$ 3270.8, observed 3271.4.

3a: [~31% yield]; ESI-MS: [MH^+] calculated for $C_{139}H_{238}N_{53}O_{37}$ 3241.8, observed 3242.4. RP-HPLC R_t = 21.90). **5**: [~26% yield]; MALDI-MS: [M+H] calculated for $C_{143}H_{245}N_{54}O_{38}$ 3326.8, observed 3326.5. RP-HPLC R_t= 23.38). **8**: [~34% yield]; MALDI-MS: [M+H] calculated for $C_{142}H_{240}N_{55}O_{40}$ 3355.8, observed 3355.9. RP-HPLC R_t= 22.48).

Supplemental Data

General materials, procedure for the synthesis of tripyrrole **6**, CD titration data for DNA binding of the tripyrrole **4**, CD spectra of hybrid **5** in the absence and presence of the target site and considering also the contribution of DNA, and EMSA titrations of **5** with T/CRE^{hs}*m*-20 and with T/CRE^{hs} in presence of excess of NON are available at http://www.chembiol.com/cgi/content/full/10/8/713/DC1.

Acknowledgments

This work was supported by the E.R.D.F., the Spanish Ministry of Science and Technology (SAF2001-3120), and the Xunta de Galicia (PGIDT00PXI20912PR and PGIDIT02BTF20901PR). J.B.B. and M.E.V. thank the Spanish Ministry of Science and Technology and the University of Santiago for their predoctoral fellowships, and M.E.V. also thanks the Human Science Frontier Foundation for a postdoctoral grant. We are very grateful to Prof. G.L. Verdine for his input and support in the early phases of this work. We also thank Prof. J. Benavente for allowing us to use radioactivity facilities and particularly Prof. M. Mosquera for his invaluable help with the data analysis.

Received: May 3, 2003
Revised: June 20, 2003
Accepted: June 23, 2003
Published: August 22, 2003

References

1. Latchman, D.S. (1998). Eukaryotic Transcription Factors (San Diego, CA: Academic Press).
2. White, R.J. (2001). Gene Transcription, Mechanism and Control (Oxford: Blackwell Science).

3. Pabo, C.O., and Sauer, R.T. (1992). Transcriptional factors: Structural families and principles of DNA recognition. Annu. Rev. Biochem. *61*, 1053–1095.
4. Burley, S.K. (1994). DNA-binding motifs from eukaryotic transcription factors. Curr. Opin. Struct. Biol. *4*, 3–11.
5. Alberts, B., Bray, D., Lewis, J., Raff, M., Roberts, K., and Watson, J.D. (1994). Control of gene expresion. In Molecular Biology of the Cell, Third Edition (New York: Garland Publishing), pp. 401–474.
6. Suzuki, M., Suckow, J., Kisters-Woike, B., Aramaki, H., and Makino, K. (1996). Multi-helical DNA-binding domains: their structures and modes of DNA-binding. Adv. Biophys. *32*, 31–52.
7. Branden, C., and Tooze, J. (1999). In Introduction to Protein Structure. (New York: Garland Publishing), pp. 129–203.
8. Latchman, D.S. (1998). In Eukaryotic Transcription Factors. (London: Academic Press), pp. 233–270.
9. Luscombe, N.M., Austin, S.E., Berman, H.M., and Thornton, J.M. (2000). An overview of the structures of protein-DNA complexes. Genome Biol. *1*, reviews001.1–001.37.
10. Jones, N. (1990). Transcriptional regulation by dimerization: two sides to an incestuous relationship. Cell *61*, 9–11.
11. Tjian, R., and Maniatis, T. (1994). Transcriptional activation: a complex puzzle with few easy pieces. Cell *77*, 5–8.
12. Wolberger, C. (1998). Combinatorial transcription factors. Curr. Opin. Genet. Dev. *8*, 552–559.
13. Khorasanizadeh, S., and Rastinejad, F. (2001). Nuclear-receptor interactions on DNA-response elements. Trends Biochem. Sci. *26*, 384–390.
14. Pedone, P.V., Ghirlando, R., Clore, G.M., Gronnenborn, A.M., Felsenfeld, G., and Omichinski, J.G. (1996). The single Cys2-His2 zinc finger domain of the GAGA protein flanked by basic residues is sufficient for high-affinity specific DNA binding. Proc. Natl. Acad. Sci. USA *93*, 2822–2826.
15. Feng, J.A., Johnson, R.C., and Dickerson, R.E. (1994). Hin recombinase bound to DNA: the origin of specificity in major and minor groove interactions. Science *263*, 348–355.
16. Ghosh, I., Yao, S., and Chmielewski, J. (1999). DNA-binding peptides. In Comprehensive Natural Products Chemistry, D. Barton and K. Nakanishi, eds. (New York: Elsevier), pp. 477–490.
17. Hu, J.C., and Sauer, R.T. (1992). The basic-region leucine-zipper family of DNA binding proteins. Nucleic Acids Mol. Biol. *6*, 82–101.
18. Hurst, H.C. (1995). Transcription factors 1: bZIP proteins. Protein Profile *2*, 101–168.
19. Weiss, M.A. (1990). Thermal unfolding studies of a leucine zipper domain and its specific DNA complex: implications for scissor's grip recognition. Biochemistry *29*, 8020–8024.
20. O'Neil, K.T., Hoess, R.H., and Degrado, W.F. (1990). Design of DNA-binding peptides based on the leucine zipper motif. Science *249*, 774–778.
21. Landschulz, W.H., Johnson, P.F., and McKnight, S.L. (1989). The DNA binding domain of the rat liver nuclear protein C/EBP is bipartite. Science *243*, 1681–1688.
22. Turner, R., and Tjian, R. (1989). Leucine repeats and an adjacent DNA binding domain mediate the formation of functional cFos-cJun heterodimers. Science *243*, 1689–1694.
23. Talanian, R.V., McKnight, C.J., and Kim, P.S. (1990). Sequence-specific DNA binding by a short peptide dimer. Science *249*, 769–771.
24. Cuenoud, B., and Shepartz, A. (1993). Altered specificity of DNA-binding proteins with transition metal dimerization domains. Science *259*, 510–513.
25. Morii, T., Shimomura, M., Morimoto, M., and Saito, I. (1993). Sequence-specific DNA binding by a geometrically constrained peptide dimer. J. Am. Chem. Soc. *115*, 1150–1151.
26. Morii, T., Saimei, Y., Okagami, M., Makino, K., and Sugiura, Y. (1997). Factors governing the sequence-selective DNA binding of geometrically constrained peptide dimers. J. Am. Chem. Soc. *119*, 3649–3655.
27. Aizawa, Y., Sugiura, Y., and Morii, T. (1999). Comparison of the sequence-selective DNA binding by peptide dimers with covalent and noncovalent dimerization domains. Biochemistry *38*, 1626–1632.
28. Caamaño, A.M., Vázquez, M.E., Martínez-Costas, J., Castedo, L., and Mascareñas, J.L. (2000). A light-modulated sequence-specific DNA-binding peptide. Angew. Chem. Int. Ed. Engl. *39*, 3104–3107.
29. Sato, S., Hagihara, M., Sugimoto, K., and Morii, T. (2002). Chemical approaches untangling sequence-specific DNA binding by proteins. Eur. J. Chem *8*, 5067–5071.
30. Zondlo, N.J., and Schepartz, A. (1999). Highly specific DNA recognition by a designed miniature protein. J. Am. Chem. Soc. *121*, 6938–6939.
31. Chin, J.W., and Schepartz, A. (2001). Concerted evolution of structure and function in a miniature protein. J. Am. Chem. Soc. *123*, 2929–2930.
32. Stanojevic, D., and Verdine, G.L. (1995). Deconstruction of GCN4/GCRE into a monomeric peptide-DNA complex. Nat. Struct. Biol. *2*, 450–457.
33. Vázquez, M.E., Caamaño, A.M., Martínez-Costas, J., Castedo, L., and Mascareñas, J.L. (2001). Design and synthesis of a peptide that binds specific DNA sequences through simultaneous interaction in the major and in the minor groove. Angew. Chem. Int. Ed. Engl. *40*, 4723–4725.
34. Ellenberger, T.E., Brandl, C.J., Struhl, K., and Harrison, S.C. (1992). The GCN4 basic region leucine zipper binds DNA as a dimer of uninterrupted α-helices: crystal structure of the protein-DNA complex. Cell *71*, 1223–1237.
35. Coll, M., Frederick, C.A., Wang, A.H.J., and Rich, A. (1987). A bifurcated hydrogen-bonded conformation in the d(A.T) base pairs of the DNA dodecamer d(CGCAAATTTGCG) and its complex with distamycin. Proc. Natl. Acad. Sci. USA *84*, 8385–8389.
36. Vázquez, M.E., Caamaño, A.M., Castedo, L., and Mascareñas, J.L. (1999). An Fmoc solid-phase approach to linear polypyrrole-peptide conjugates. Tetrahedron Lett. *40*, 3621–3624.
37. He, G.-X., Browne, K.A., Groppe, J.C., Blaskó, A., Mei, H.-Y., and Bruice, T.C. (1993). Chemistry of phosphodiesters, DNA and models. 3. Microgonotropens and their interactions with DNA. 1. Synthesis of the tripyrrole peptides dien-microgonotropen-a, -b, and -c and characterization of their interactions with dsDNA. J. Am. Chem. Soc. *115*, 7061–7071.
38. Rentzeperis, D., Marky, L.A., Dwyer, T.J., Geierstanger, B.H., Pelton, J.G., and Wemmer, D.A. (1995). Interaction of minor groove ligands to an AAATT/AATTT site: correlation of thermodynamic characterization and solution structure. Biochemistry *34*, 2937–2945.
39. Chen, F.-M., and Sha, F. (1998). Circular dichroic and kinetic differentiation of DNA binding modes of distamycin. Biochemistry *37*, 11143–11151.
40. Satz, A.L., and Bruice, T.C. (2002). Recognition in the minor groove of double-stranded DNA by microgonotropens. Acc. Chem. Res. *35*, 86–95.
41. Dervan, P. (2001). Molecular recognition of DNA by small molecules. Bioorg. Med. Chem. *9*, 2215–2235.
42. Hollenbeck, J.J., and Oakley, M.G. (2000). GCN4 binds with high affinity to DNA sequences containing a single consensus half-site. Biochemistry *39*, 6380–6389.
43. Mammen, M., Choi, S.-K., and Whitesides, G.M. (1998). Polyvalent interactions in biological systems: implications for design and use of multivalent ligands and inhibitors. Angew. Chem. Int. Ed. Engl. *37*, 2755–2794.
44. Mammen, M., Shakhnovich, E.I., and Whitesides, G.M. (1998). Using a convenient, quantitative model for torsional entropy to establish qualitative trends for molecular processes that restrict conformational freedom. J. Org. Chem. *63*, 3168–3175.